CODIB-D-79
CRAG 2-61

CENTRAL INTELLIGENCE AGENCY

OFFICE OF CENTRAL REFERENCE

14 March 1961

MEMORANDUM FOR: Committee for Documentation
Central Reference Advisory Committee

SUBJECT: The Office of Central Reference (OCR) Method
of Evaluation of Systems Proposals

The attached paper, circulated for information, outlines the factors

considered by OCR during the course of evaluating proposals affecting

our information handling systems.

STAT

Paul A. Borel
Assistant Director
Central Reference

Attachment

OENTRAL INTELLIGENCE AGENCY

OFFICE OF CENTRAL REFERENCE

The Office of Central Reference (OCR) Method of
Evaluation of Systems Proposals

21 February 1961

Introduction:

To cope with its information handling problems, the Central Intelligence

Agency has over the past 13 years developed an information processing center

which comprehensively indexes and stores that information which is collected

and, as a service of common concern, renders daily support to analysts at work

in all parts of the U. S. Government's intelligence community.

The Office of Central Reference fulfills its mission--primarily that of

document and information retrieval rather than data processing--through a

complex of interlocking and interdependent, manipulative and non-manipulative

indexing sub-systems, most of which rely heavily on commercially available

punch card equipment (EAM) for implementation. Each sub-system serves a

specific function and differs in scope and form, yet each must meet certain

theoretical and practical standards which have been established for the system

as a whole. The number of sub-systems is kept to a minimum and a conscious

effort is made to sustain and improve the efficiency of those which are necessary.

Consumer demand is balanced against availability of manpower and machine time

prior to the establishment of a new sub-system within a given OCR Division.

Selected documents are processed into that sub-system which, when related to

consumer need for depth of indexing and manipulation of indexed data represents

the most economical use of time and equipment.

The Automation Development Group, OCR, participates in the preparation

of staff studies for the AD'CR which set forth conclusive comparisons of the

effectiveness and cost of operating a particular OCR system of information

processing and retrieval with that of proposed systems (solicited and unsolicited).

General Criteria:

Not every systems proposal received by the AD/CR results in the preparation

of a staff study. Proposals which are similar to previously evaluated and rejected

proposals, and those which do not retain all the substantive qualities which

characterize the present system require only the preparation of a memorandum

for the record approved by the DAD'CR. Proposals which encompass processing

techniques, output options and 'or use of equipment different from but of interest

to OCR are carefully evaluated and the results presented to the AD/CR in the form

of staff studies. This activity is an integral part of the OCR program of re-appraisal.

Identification of problem areas with a view to improving the services within the

present frame work includes the adoption of that new system which demonstrates

material advantages over the current system. Identification of consumer requirements

and development of technique for ensuring fulfilment of these requirements are of

more concern to OCR than selection of equipment per se.

Once determined that a staff study is in order, the Automation Development

Staff reads the proposal and (1) identifies those sub-systems within the OCR Complex

which would be affected by the adoption of the proposed system. (These sub-systems

are used as a base from which comparisons relative to procedures, manpower

requirements, time, space, and equipment as well as to the qualities of the outputs

from the systems are formulated.); (2) recommends that designated personnel

from various sub-systems work with the ADG in the preparation of a preliminary

study for the DAD/CR who in turn participates in synthesizing the elements of the

staff study for the AD/CR; (3) requests from the appropriate Division Chiefs those

existing written records and statistical tables necessary to the preparation of the

required staff study. These may include: Statement of mission; service

responsibilities, philosophy and capabilities; description of document holdings;

receipt, records and dissemination procedures; intellectual processing techniques

(dossier building techniques, machine support for report writing and publications

preparation, indexing for document and information retrieval, special projects

for data handling); description of coding schemes, machine supported working

aids for indexers; creation, maintenance and servicing of files (indexes and

document) including machine support; identification of equipment used; statement

of output options; personnel requirements at various phases of processing cycle;

identification of strengths and problem areas of current system; statistical tables

setting forth personnel, space, time, equipment, volume requirements of the current

system; and (4) coordinates the efforts of personnel involved in the evaluation of the

proposal and drafts the preliminary study.

Preliminary steps necessary to the preparation of a systems staff study can be summarized as follows:

1. Determine the points of similarity or identity between the present system and the proposed system (at each stage of processing cycle).

2. Identify the first point in the processing cycle which would be affected by adoption of the proposed system.

3. Summarize the changes within the present system which would be necessitated by the adoption of the proposed system (at each stage of processing cycle).

4. Set forth the advantages of the present system over the proposed system.

5. Set forth the advantages of the proposed system over the present system.

6. Make comparative tables of cost factors (by stage of processing cycle) showing requirements of the two systems for manpower, space, time and equipment.

Specific Criteria:

As an integral part of the systems proposal evaluation process, the Automation Development Group is accumulating a list of check-points which are used during the preparation of the preliminary study for the DAD/CR. Grouped according to the stages of processing within the present OCR system there exists a certain amount of redundancy between sections of this list. Not all these check-points are applicable to every proposal.

I. Physical Handling of Documents

1. Will adoption of proposed system necessitate changes in the established practices for receipt, control and dissemination of documents?

2. Will necessary controls be more automated? Can indexing and dissemination be combined? Can dissemination be automated in an effective manner?

3. Can overlapping reading panels be discontinued?

4. Will requirements for recovery of documents still in processing be discontinued?

II. Intellectual Problems

1. Will the proposed system give stronger support to the dossier building and report preparation efforts of the OCR Divisions?

2. Does the proposed system provide for data processing (manipulation of quantifiable information extracted from documents and formatted for input to the machine system)? To what extent must data be ordered before input and at what cost? To what extent does proposed system provide for recovery of relevant material, reduce redundancy and prevent false drops? Will equipment accept and manipulate both alphabetic and numeric information? Variable lengthwords, fixed and open field words, variable length unit records, linkage symbols? Must the indexer identify these elements as well as separate related and independent variables?

3. Will the extraction of manipulative, formattable, and quantifiable information be performed by the same group of people as the coding for control of concepts contained within the same documents? (data handling vs document retrieval)

4. Will the type of coding for input to the proposed system be different from that which is performed for the present system? Will the present indexing tools (code schemes and machine supported working aids) be applicable in the proposed system?

5. What effect will the adoption of the proposed system have on the present indexing procedures, transcription procedures, training of indexers?

6. Does the proposed system provide for levels of input, levels of output, multiple end product and manipulation of data contained in the index?

7. Will indexing for the proposed system be easier than or more difficult than that done for the present system? Will the indexer be able to expect more and faster machine support for his own job of ordering the information? Will dictionary building be easier to perform with new equipment or will EAM equipment still be used for this support function?

Can the dictionary entries which control the ordering of information relative to concepts, commodities, organizations, areas, and persons be as a single file or will it still be necessary to categorize at the point of input?

8. Will specialization on the part of the input officer be possible or desirable within the new system? Will input and retrieval be performed by the same group of persons?

9. Will the problem of establishing "indexing depths" continue to exist in the proposed system? Will the input officer receive any help from the machines in identifying and establishing indexing depths?

10. Automatic indexing and abstracting devices are on the horizon. Does the proposed system lend itself to the use of these devices without making necessary the addition of steps in the processing system?

11. How will the proposed system assist the document officer in the preparation of abstracts necessary to the present OCR service? Will the preparation of these bibliographic tapes be made easier by adoption of proposed system?

12. Will coding of information be on visible record, capable of being verified prior to and after acceptance by the machine and will correction, deletion and updating of information entered into the machine record be easy to accomplish?

III. Creation, Maintenance and Servicing of Machine Indexes

1. Does proposed system advocate one universal arrangement of all inputs as opposed to a series of sub-systems for special purposes? Will separate indexes be maintained for the control of photographic, organization, biographic, subject, commodity, and area information? Will separate files be maintained for each formatted file necessary to data processing activities?

2. Will the proposed systems indexes be self-organizing?

3. Will manual search be possible? Will all service be necessarily tied to the use of machines?

4. How will updating and corrections be accomplished? What will be the timing for updating and corrections and how will this be reflected in service?

5. Identify the output options: citations, abstracts, aperture cards, hard copy, machine printed listings, other.

6. Will it be possible to receive: print-outs of the entire index record in its natural or specified type of entry; entries selected out of the record and printed according to a designated format; a print-out of citations; a combination of above types of print-outs; a print-out of the results of the statistical manipulation of specified unit records or specified portions of unit records?

7. Will the request format and language of the proposed system differ from that used in the present system? Will the system accept standing, routine, priority, priority override, multiple search requests? Will search questions be limited by number of question words? Memory capacity?

8. What economies will be effected by the adoption of the proposed system for the creation, maintenance and servicing of the machine indexes?

9. What qualitative advantages are inherent in the end products offered by the proposed system? Within a record? Overall?

10. Is there a demonstrated need for all end products offered by the proposed system?

IV. Establishment, Maintenance and Servicing of Document Holdings

1. How will the structure of the document holdings for the proposed system differ from that of the present system? Will Bell and Howell reductions, aperture cards, Actifilm, roll film, and hard copy be included?

2. Will the proposed system involve the reduction of a greater percentage of the document holdings to aperture card? Or similar form?

3. Will preparation, maintenance and servicing of the document files be further automated by the proposed system?

4. Does the proposed system make more difficult the preparation of expendable hard copy prints for the consumer?

5. Does the proposed system provide for the manual access to aperture cards?

6. What effect would adoption of proposed system have on the manpower, space, time for servicing a given request and quality of end product based on document holdings?

7. Output options for the present system include: Bell and Howell card tape for first page or one page documents; aperture cards or actifilm reproductions; print from aperture cards, actifilm or roll film, reproductions of hard copy documents; loan of hard copy document per se; reproduction or loan of maps, photos, charts through Graphics Register; use of dossiers for biographic and industrial information within the Registers. Will output options for proposed system differ in format? In timing? Will they be more easily produced and involve the use of less manpower?

8. Does the proposed system provide for hard copy back-up for formatted files used in data processing as well as for the unformatted files created for document recovery?

V. Equipment (Other than EAM)

1. On what equipment is the Proposed System dependent for implementation? Identify that equipment necessary to the central system; supporting equipment; filming and reproduction equipment necessary to production of proposed output options (Machine indexes and document holdings). Identify storage media. If magnetic tape, is entry based on BCD or octal code?

2. Is the equipment modular? Can the system be modified by addition without requiring replacement of central system equipment? After amortization of initial cost would equipment proposed for system be likely to have become obsolete?

3. Will proposed system require the use of manual or semi-automatic equipment to any degree? How much of the presently used EAM equipment could OCR dispose of if proposed system were adopted?

4. What protection does the new equipment offer against loss of stored information?

5. What backup in the new system provides for service in event of automatic input equipment failure?

6. Is the proposal fitted to the use of equipment not requiring further research and development? Is it commercially available on rental basis? Purchase basis? Proposed amortization?

7. Can the equipment be readily understood and operated? As user ability and need increases, can new equipment be added to the system?

8. Highly complex equipment and procedures may be necessary when a highly complex or exacting job must be done, but need should be allowed to arise first. There is nothing to be gained from adding degrees of freedom to a system until it is at least assured that the users can understand and master these modifications. Does OCR mission justify implementation of Proposed System (processing techniques and equipment)?

9. Does equipment possess searching logic not required for the processing and retrieval of information?

10. Will adoption of system based on new equipment require the continuing support of contractor personnel? Will use of contractor personnel increase or reduce annual personnel costs?

11. What effect will adoption of proposed equipment have on present Machine Division Staff?

12. How difficult will corrections to existing files be with new equipment?

13. Can equipment be used for the purpose of dictionary building? How much support can the input officer expect from machines used for implementing the proposed system?

14. What is the initial cost of the equipment; the spare parts requirements; and the estimated cost of preventative maintenance?

15. Would equipment (including input and retrieval techniques) be tested in an operating environment prior to final adoption (as in case of minicard equipment)? Estimate cost of test? For what period of time would OCR be required to operate its present system should the proposed system be adopted?

VI. Space Requirements

Compare the space requirements of the present system with that of the proposed system in terms of personnel, equipment, index storage, document storage, retirement space for index and document files.

VII. Conversion Problems

1. Would it be possible to convert the present OCR holdings (IBM card files) to the media for processing with the equipment on which the proposed system is based? Will present holdings need be maintained and serviced in present form?

2. Would the conversion involve physical problems only? Would it be feasible to attempt intellectual conversion at the same time?

3. Identify the cost of conversion in terms of manpower, time for completion and interruption of service. Would cost of conversion approximate the cost of original coding? Would resultant product be as effective?

4. Even though converted to new media would present holdings need to be maintained and serviced as separate files within the new system?

VIII. Human Factors

1. Will the over-all manpower requirements be lowered or increased by adoption of the proposed system?

2. Will the present staff need to be retrained in most of their functions in order to perform within the proposed system?

3. Will adoption of the proposed system raise or lower the intellectual requirements placed on the personnel engaged in the operation of the system?

4. Will experience to date within the present system assist in the training of personnel to operate within the proposed system?

5. Will the "need to know", relative to the total system be a must for the document analyst, machines personnel and reference personnel as is the case in the present system?

6. Must the consumer take a more active part in the creation of inputs for the proposed system? Will his participation be supplemented by or standardized by a central indexing group?

7. Will adoption of the proposed system advance the cause of compatibility within the intelligence community? Theoretically or realistically?

- 11 -

IX. Flexibility

1. Is the proposed system more flexible than the present system? Will it meet the changing functional requirements of the OCR complex? Is it modular?

2. What are the speeds at all facets of the operation of any given sub-system? Identify the limitations and capabilities of the system at all stages of the processing cycle. How do these compare with the limitations and capabilities of the present system? Are identified weaknesses within the present system non-existant in proposed system? Does the proposed system retain all the identified musts of the present system?

3. How automatic is the proposed system? What manual operations are necessary to the functioning of the system as a whole? How do these compare to those necessary to the present system?

4. Is proposed system adequate to perform the complex information handling task which faces the Office of Central Reference within the limits of an established budget and with concomitant improvements in service.

The foregoing checklist represents the present criteria used by the ADG for evaluating systems proposals. As each proposal is considered this list is up-dated to reflect new decisions reached with regard to techniques or equipment. Any proposal which is confined to generalities and lacking in sufficient detail to permit tabulated comparison precludes the possibility of preparing a conclusive staff study. However, proposals which seem promising despite insufficient detail may be re-submitted.